

Fig. 1

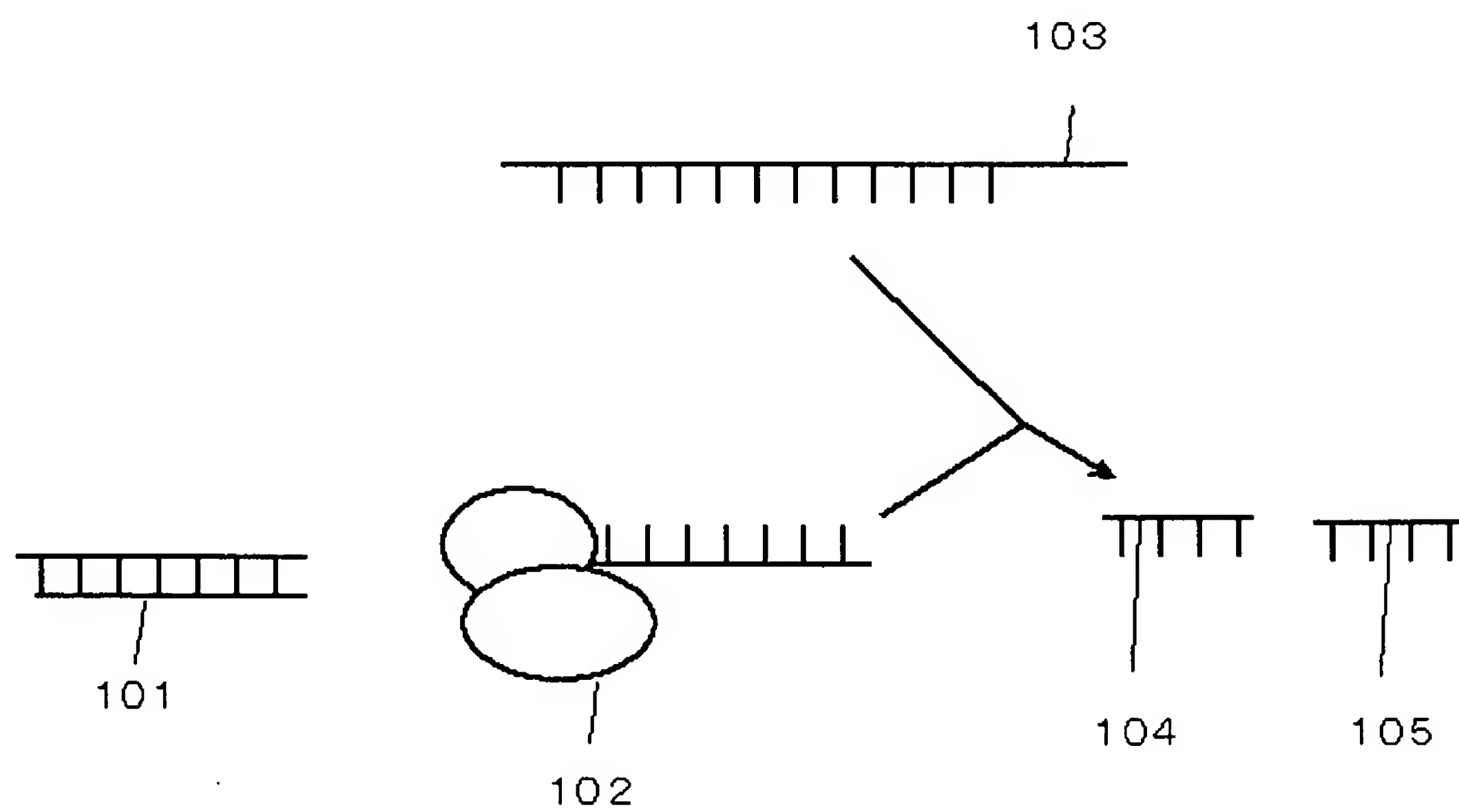


Fig. 2

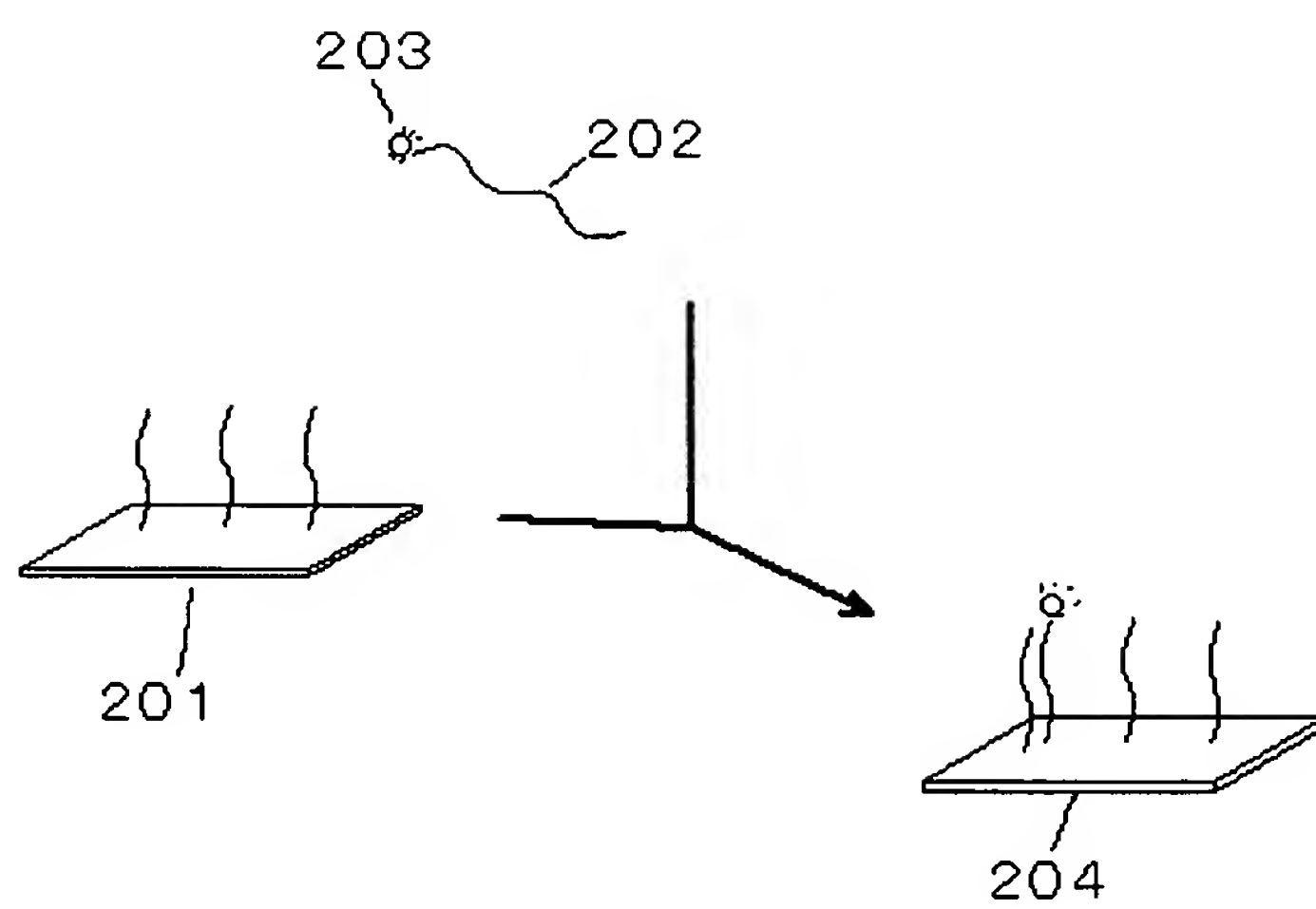


Fig 3

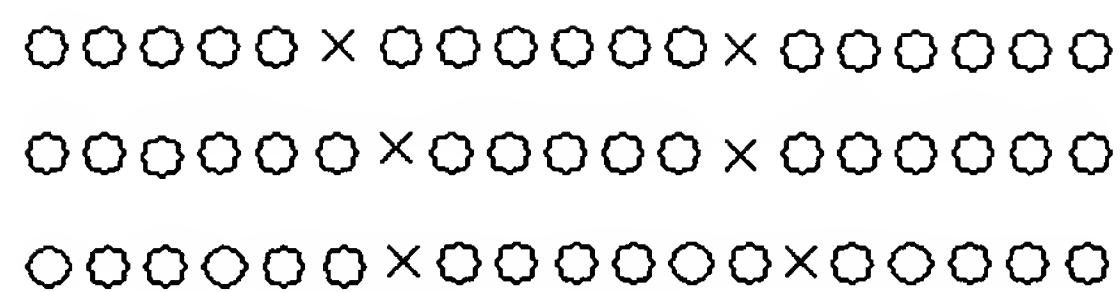


Fig. 4

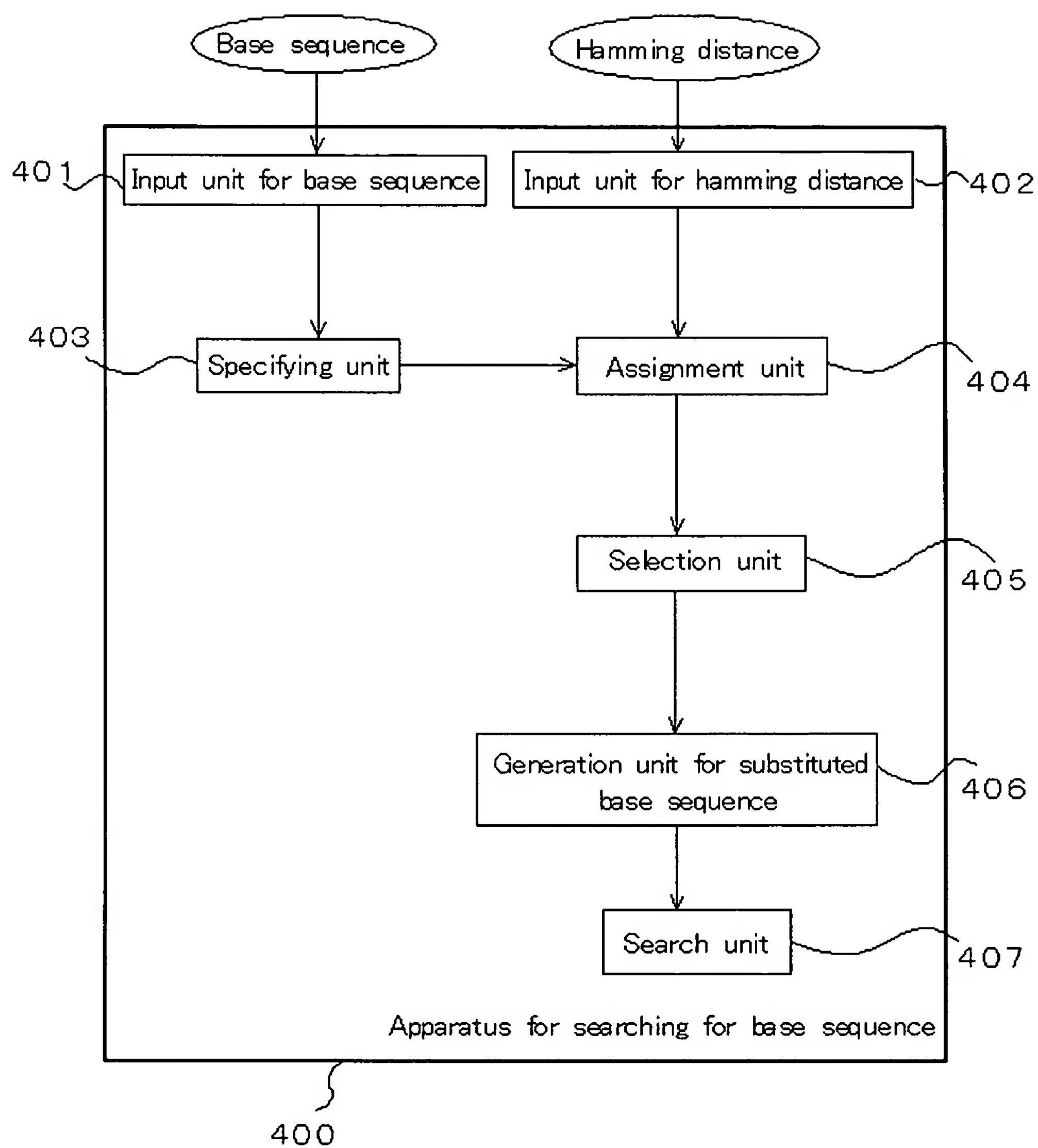


Fig. 5

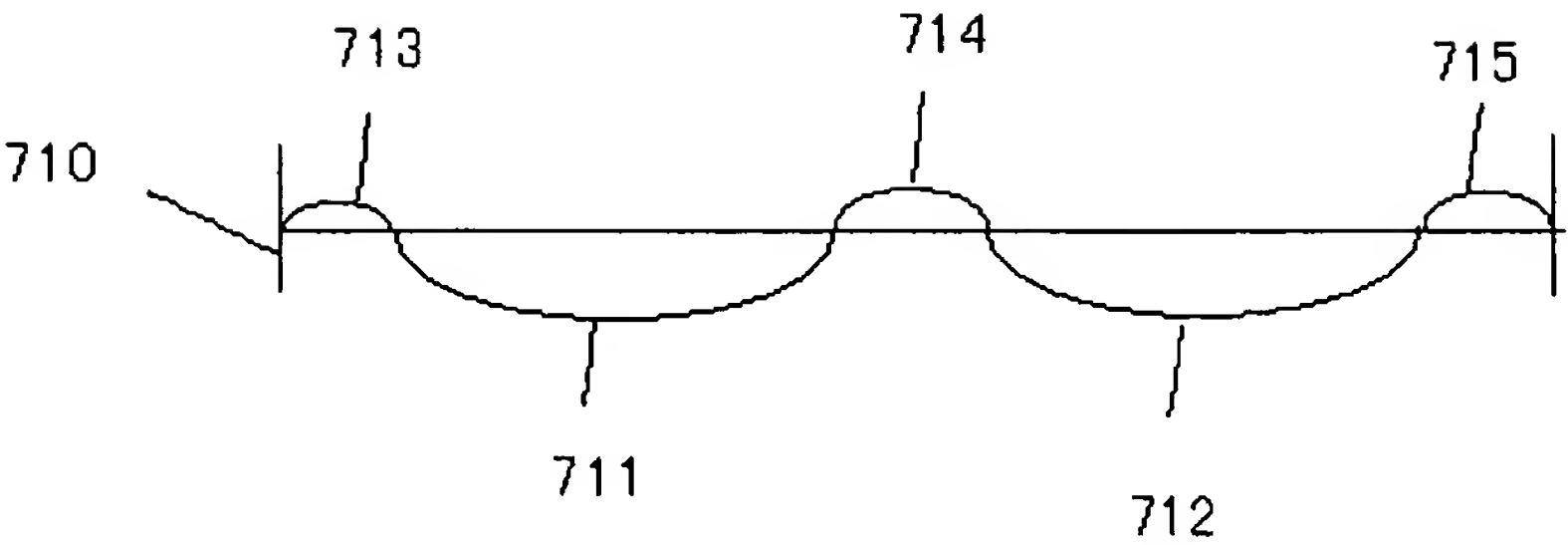
A	C	C	G	T	A	G	T	A	T	A	T	G	C	T	G	C	T	G
				X						X					X			
A	C	C	G	C	A	G	T	A	T	C	T	G	C	T	T	C	T	G

Fig. 6

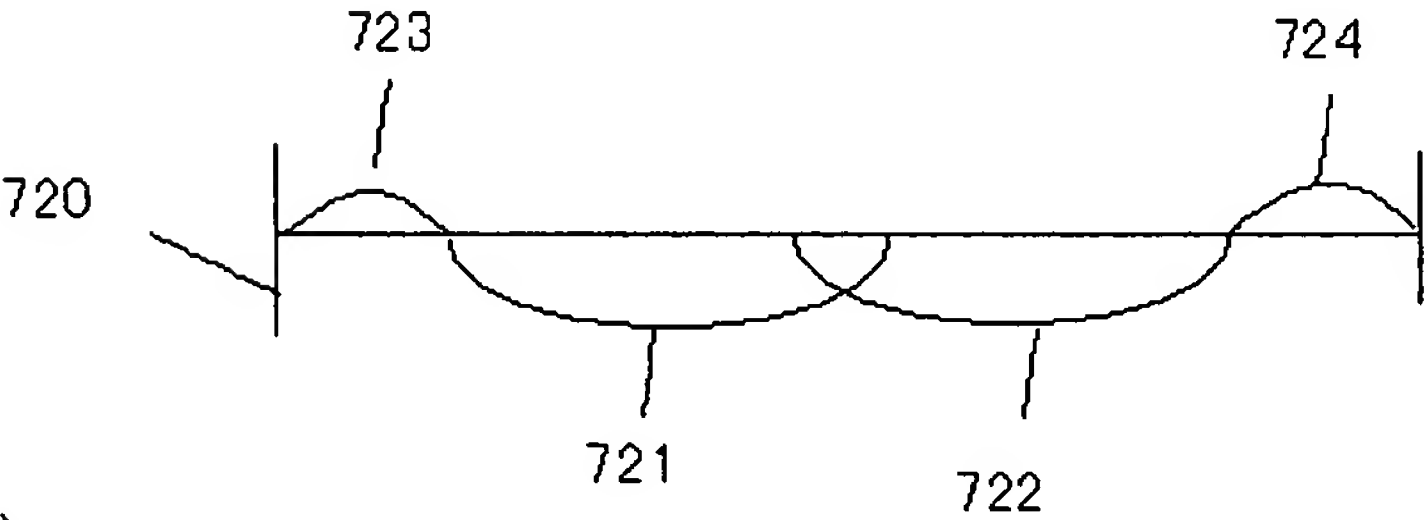
$$d_H(S,T) = |\{i \mid S_i \neq T_i, i=1,2,\dots,n\}|$$

Fig 7

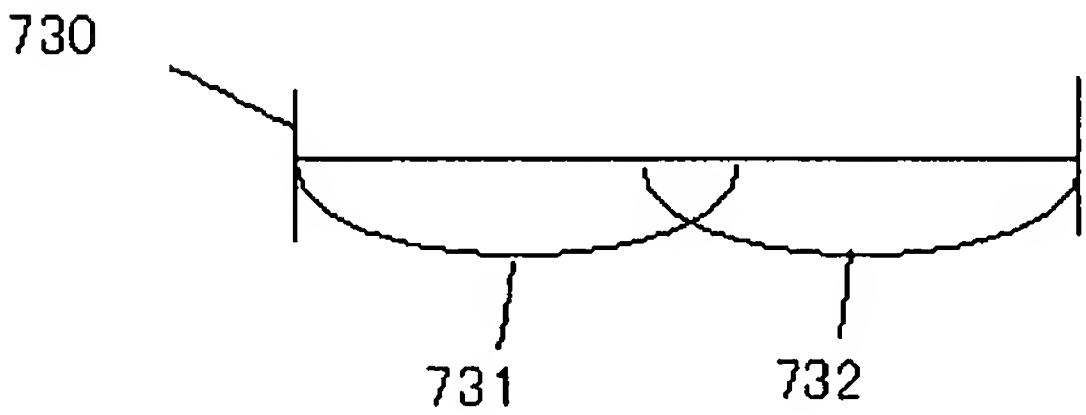
(1)



(2)



(3)



(4)

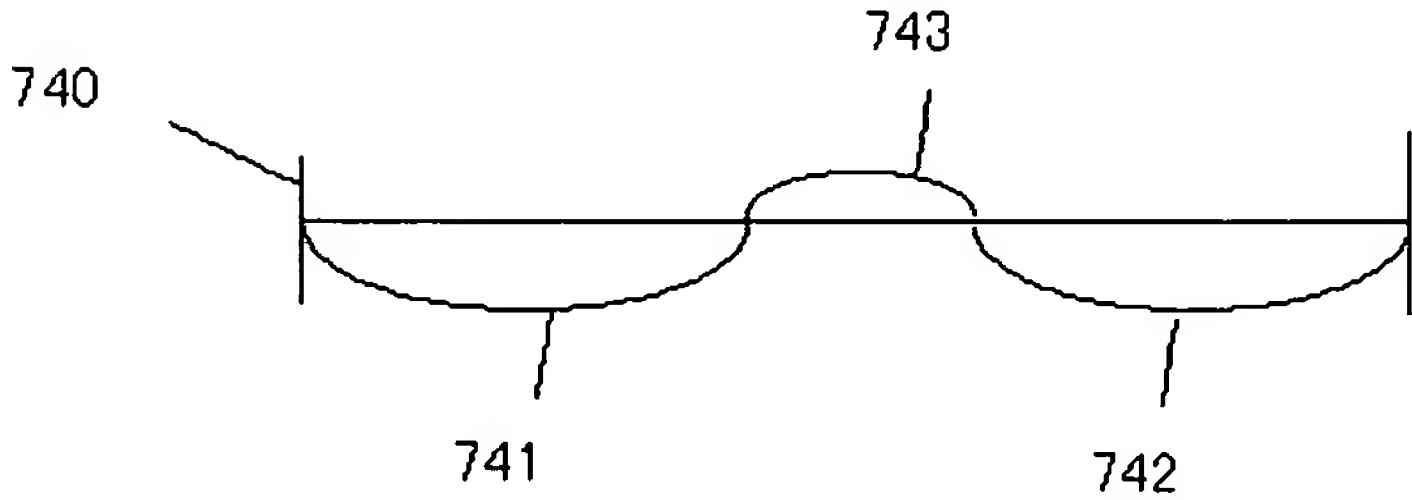


Fig 8

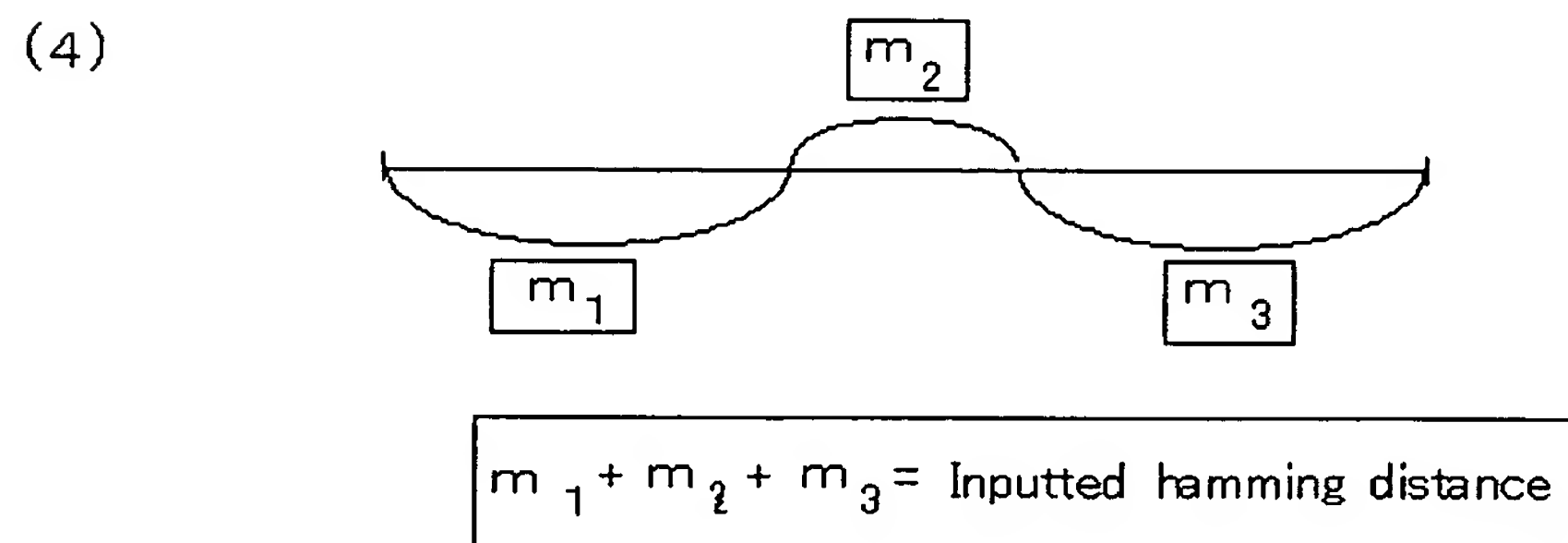
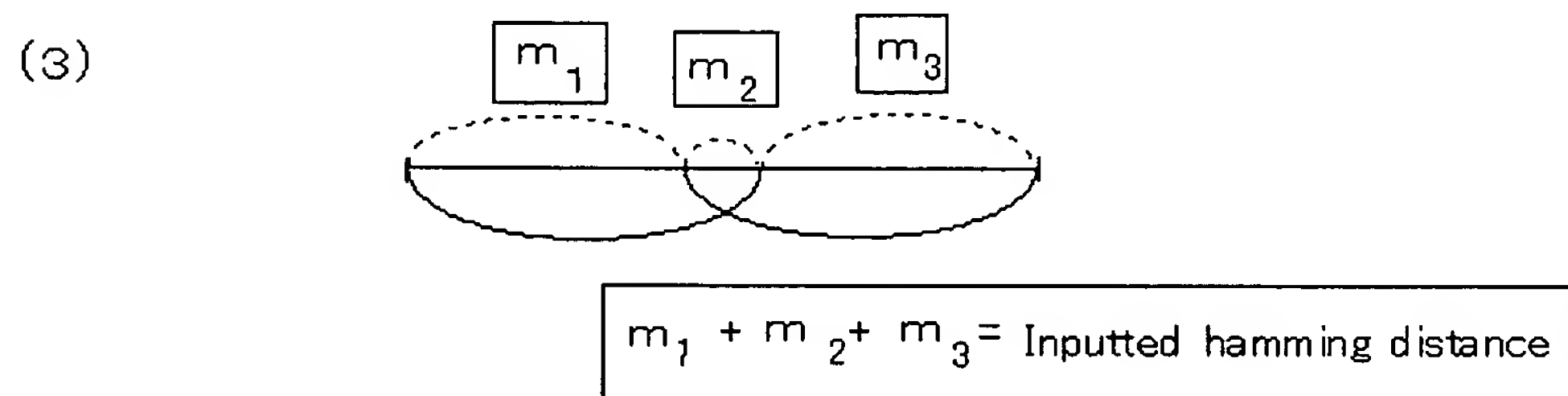
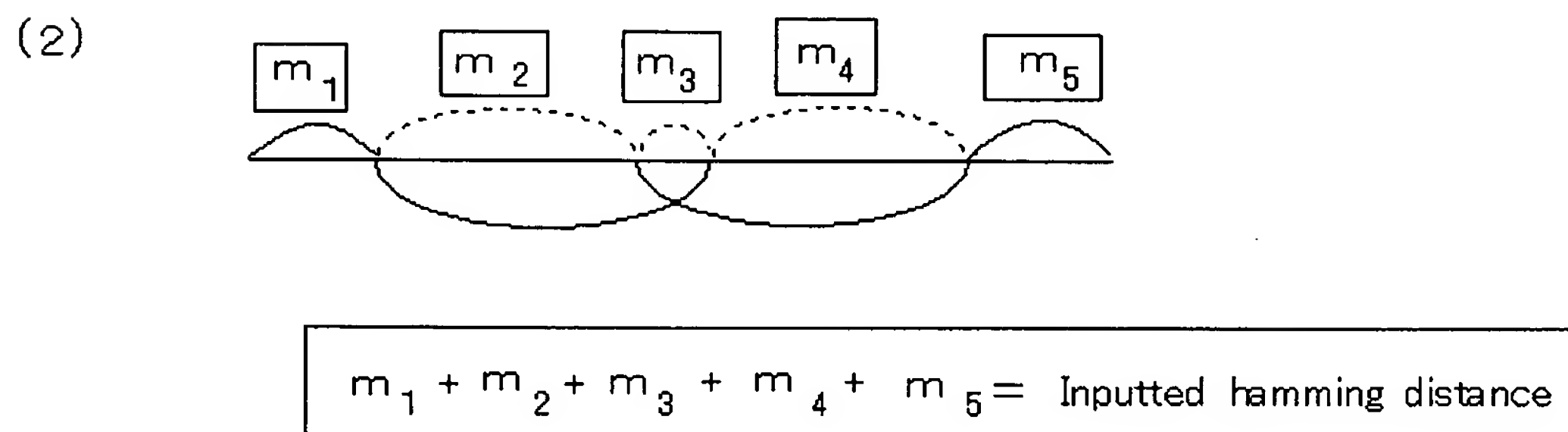
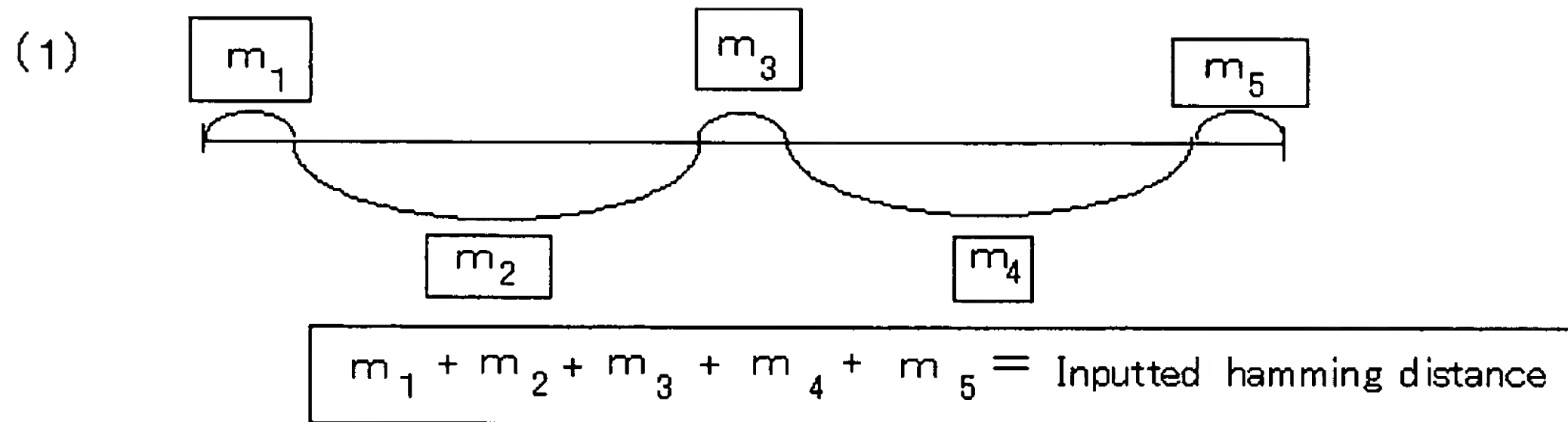


Fig 9

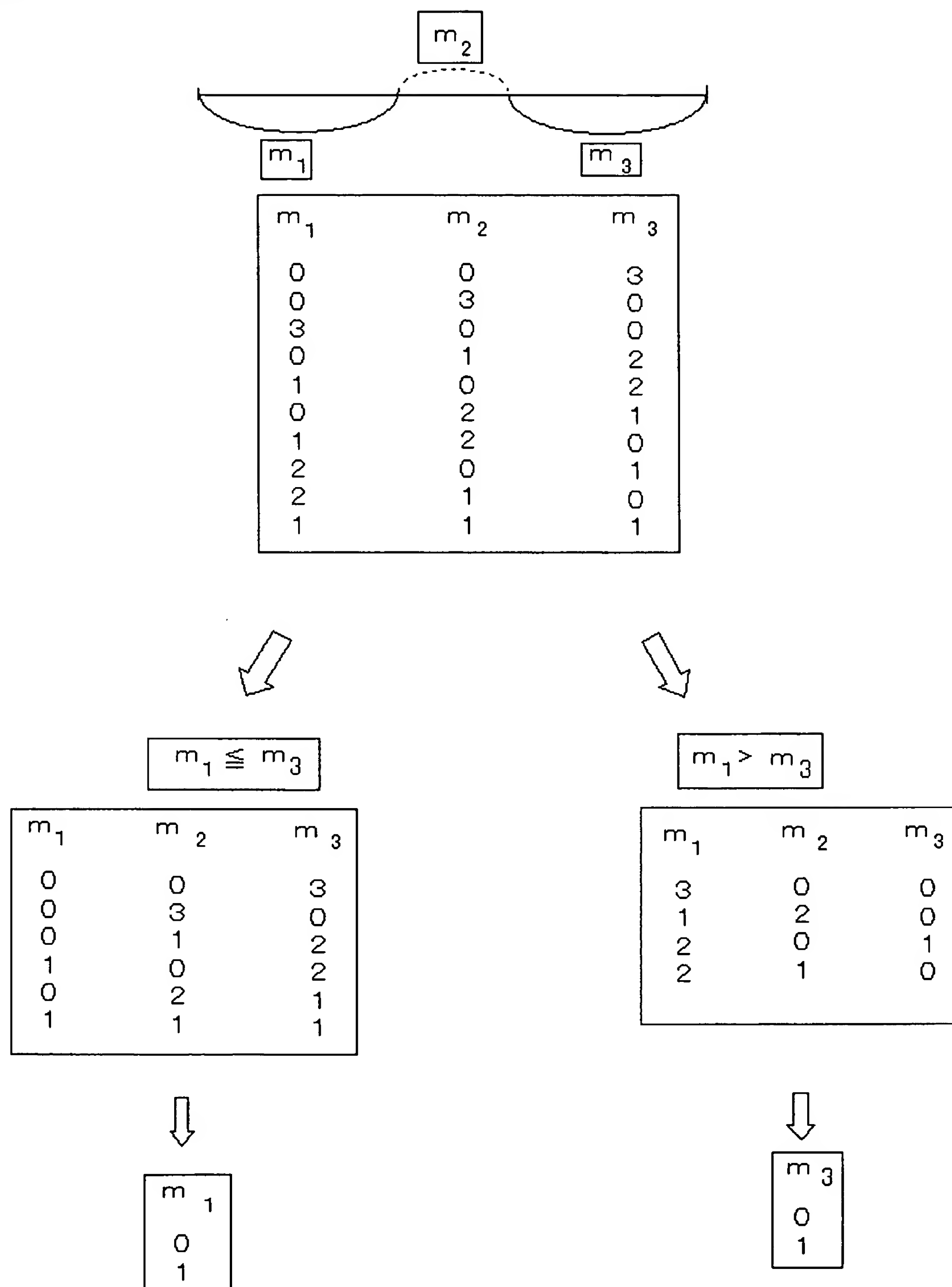


Fig 10

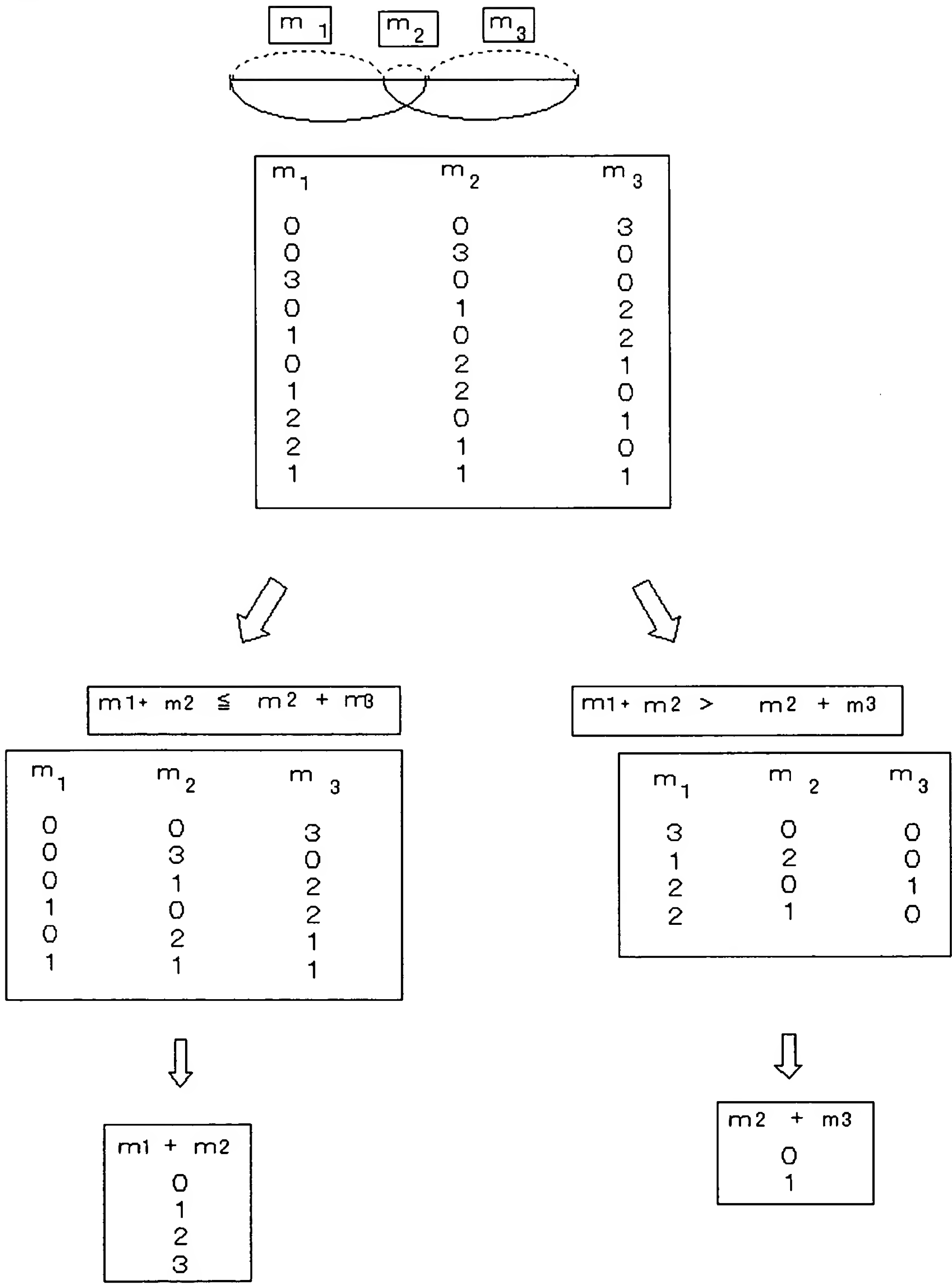


Fig. 11

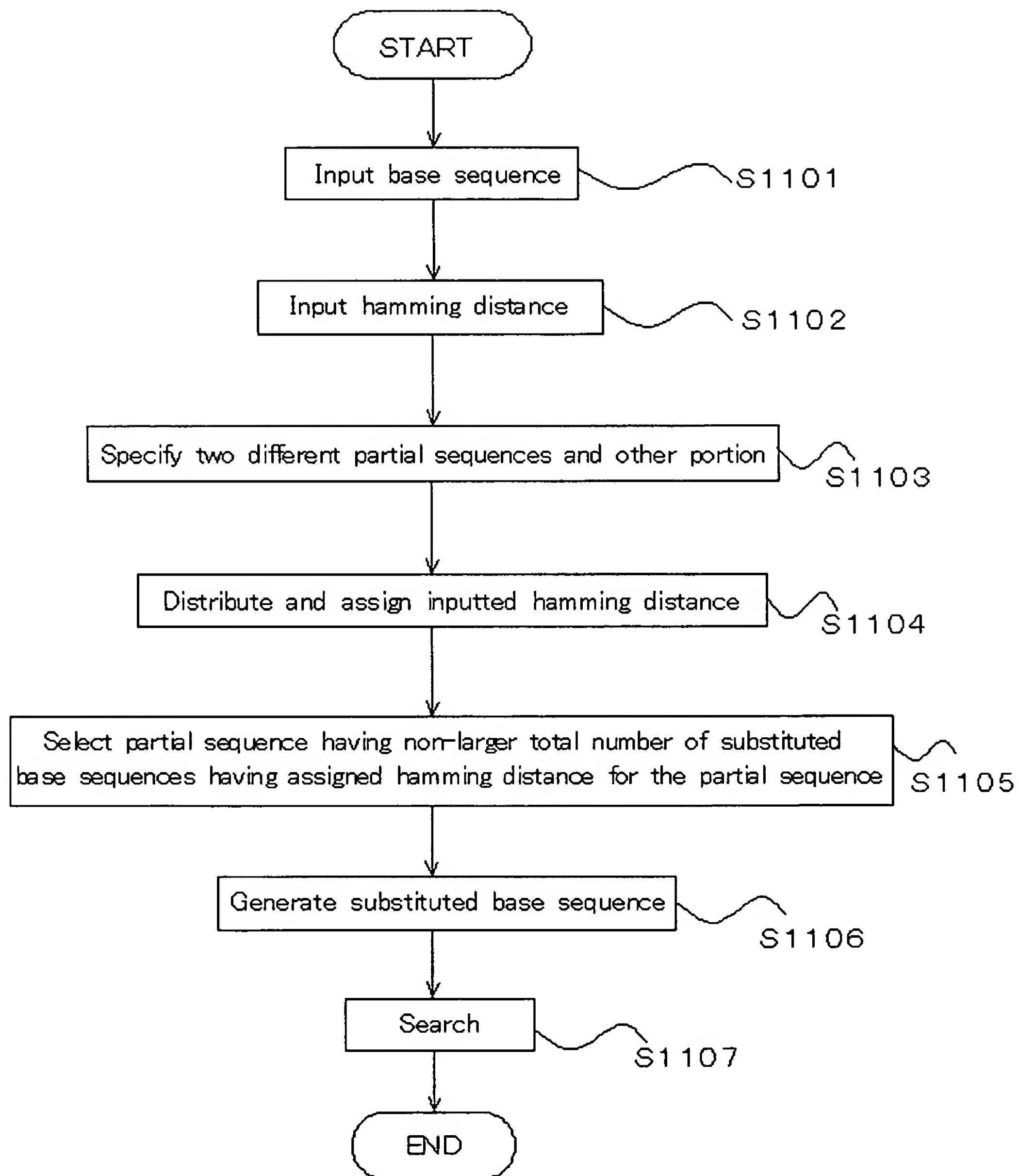


Fig. 12

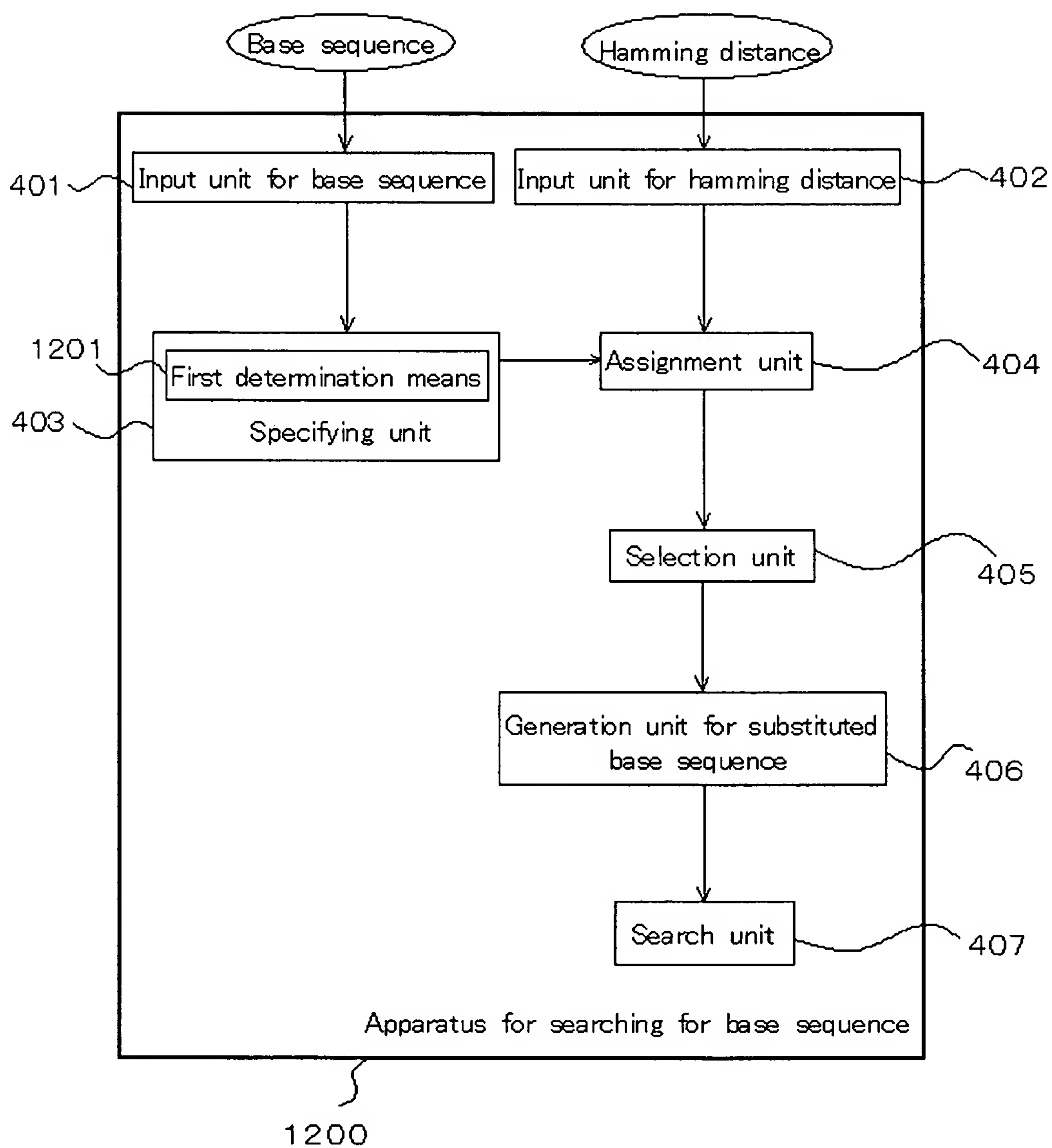


Fig 13

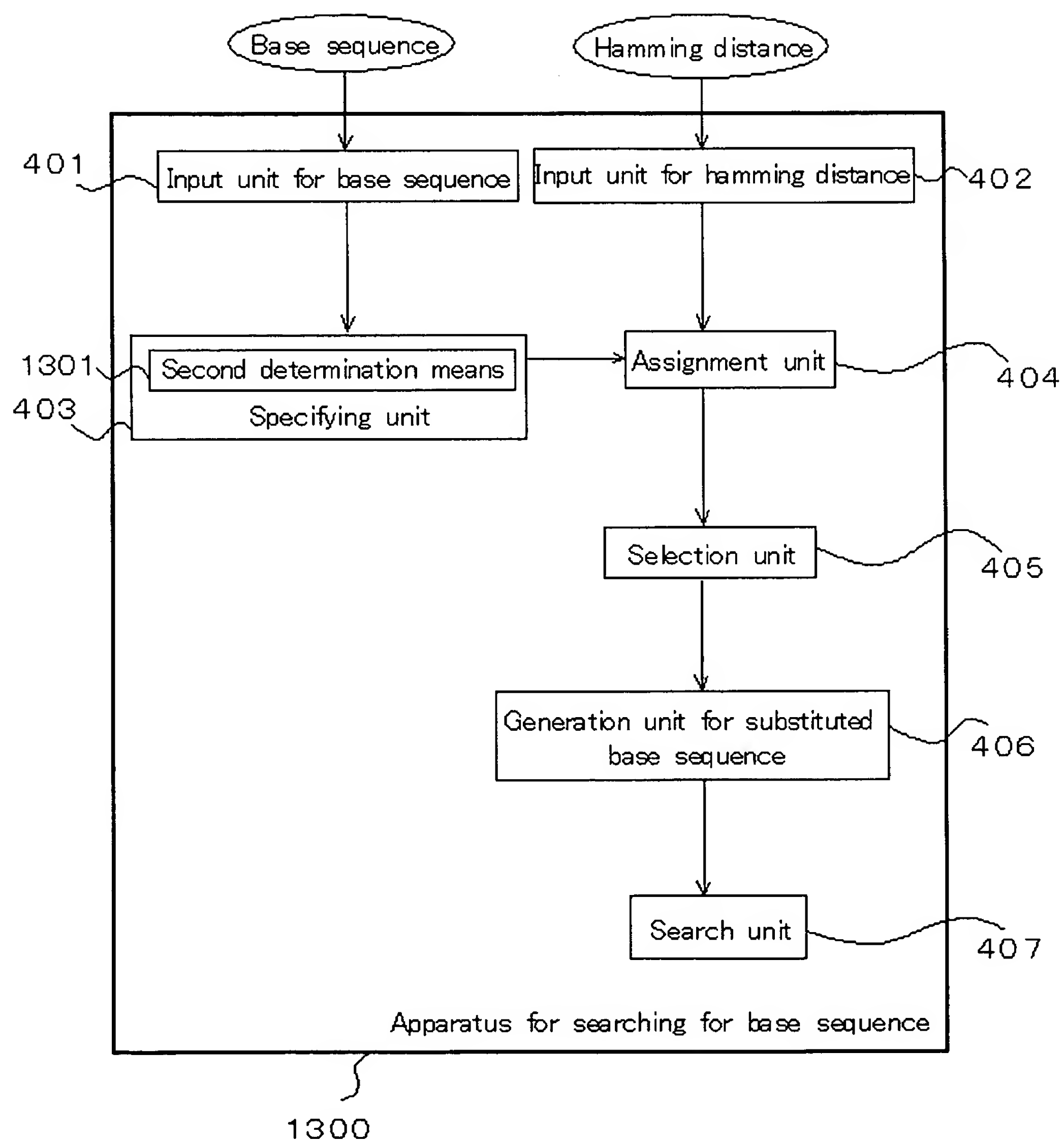


Fig. 14

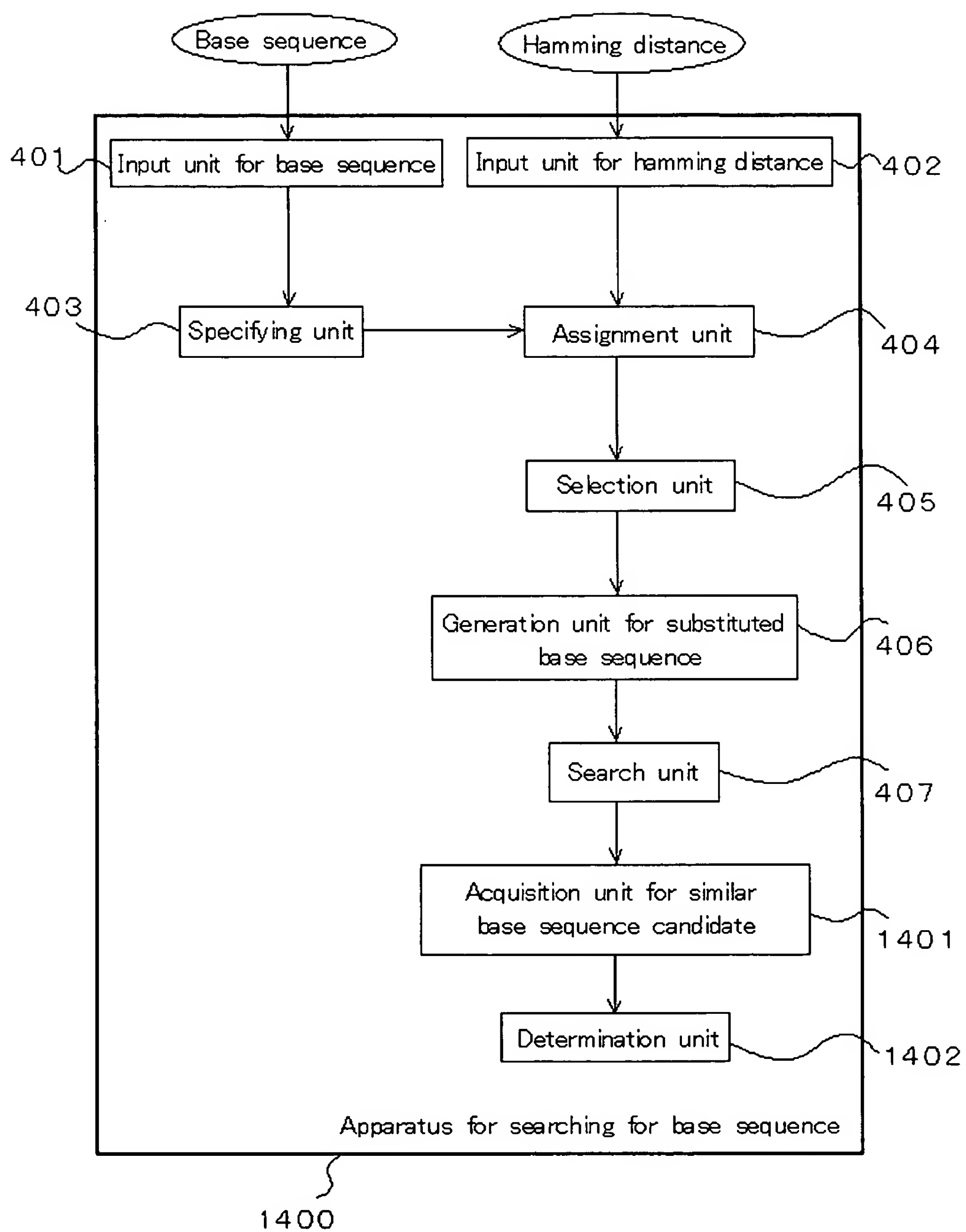


Fig 15

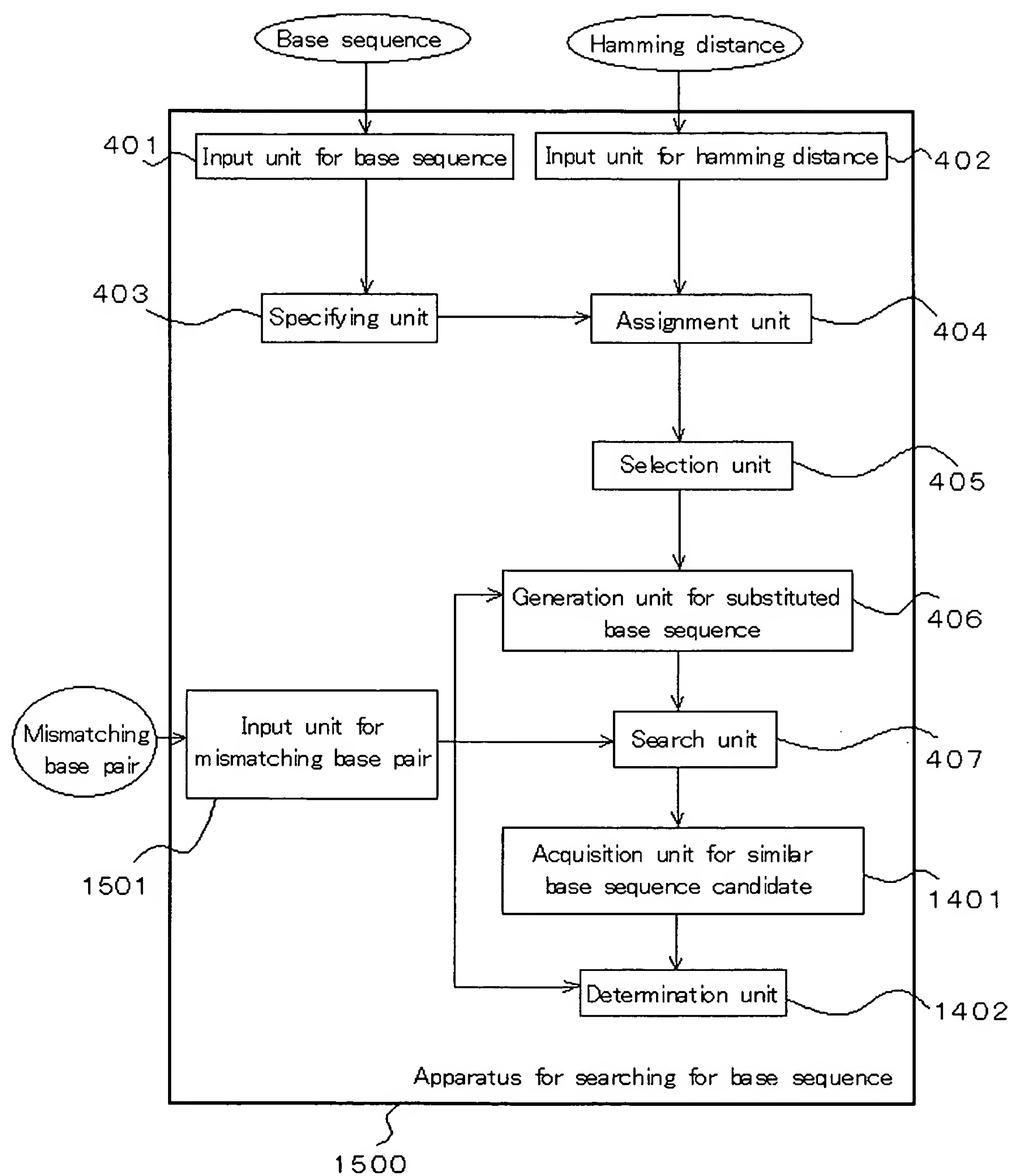


Fig. 16

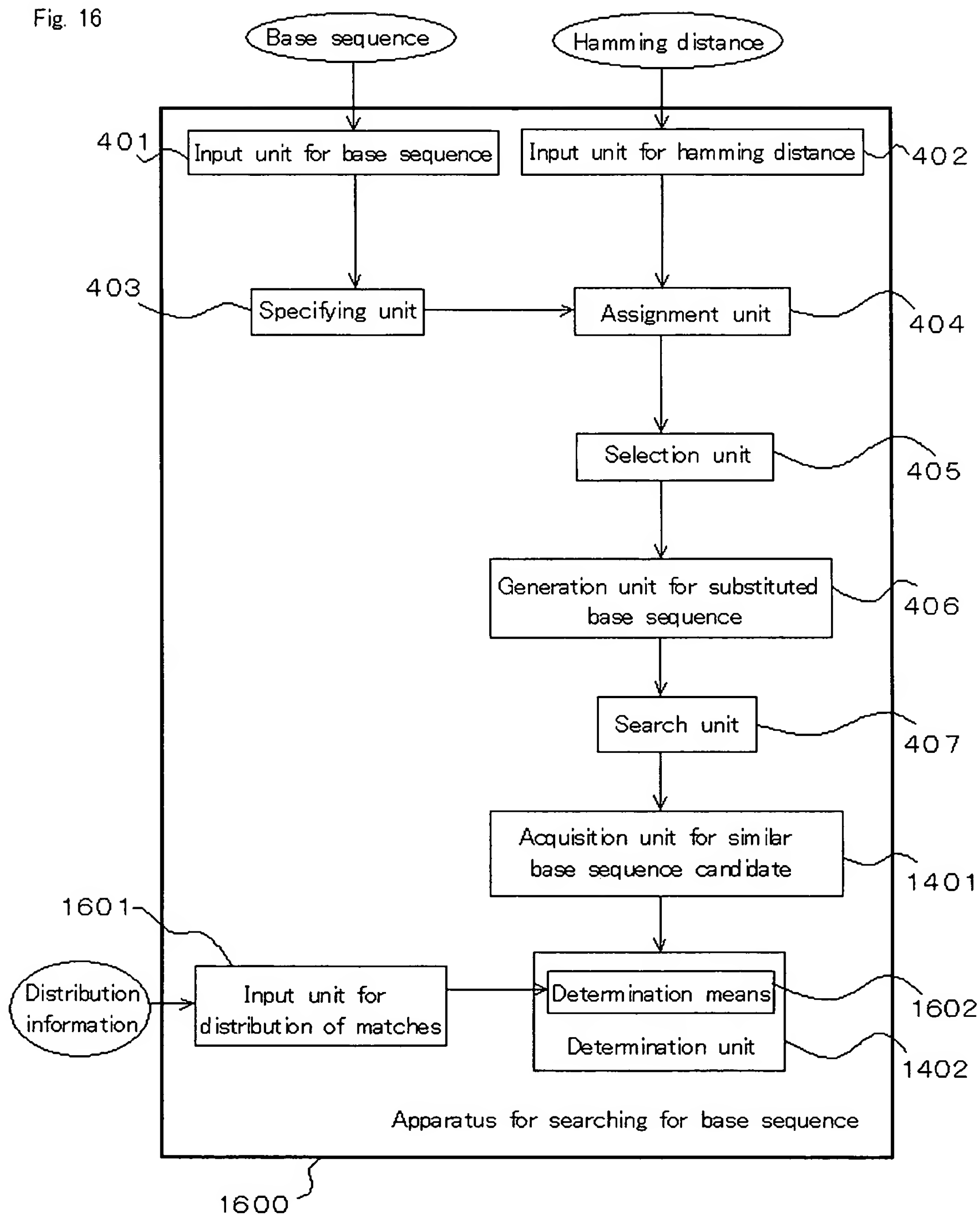


Fig 17

```

distributeHammingDistance(int P, int H, int nSize, int* vec)
{
    int h;

    if (P==1) {
        vec[1] = h;

        /*
         * one of the assignments of the hamming distance to all partial
         * sequences is completed, so that the hamming distances
         * stored in vec are outputted
         */
        for (int i = 1; i <= nSize; i = i + 1) {
            printf("Part %d th: %d", i, vec[i]);
            /* a separator or a terminator is outputted */
            if (i != nSize) {
                /* a comma is outputted as a separator */
                printf(", ");
            }
            else {
                /* a break is outputted as a terminator */
                printf("\n");
            }
        }
    }
    else {
        for (h = 0; h <= H; h = h + 1) {
            vec[P] = h;
            distributeHammingDistance(P - 1, H - h, nSize, vec);
        }
    }
}

```

Fig 18

```
for (I1 = 0; I1 < L; I1 = I1 + 1) {  
  for (I2 = I1 + 1; I2 < L; I2 = I2 + 1) {  
    foreach a1 in {A, C, G, T} {  
      if (S[I1] != a1 ) {  
        foreach a2 in {A, C, G, T} {  
          if (S[I2] != a2) {  
            Generate the substituted base sequence  
            by substituting I1-th base of S to a1,  
            and the I2-th base to a2  
          }  
        }  
      }  
    }  
  }  
}
```

Fig 19

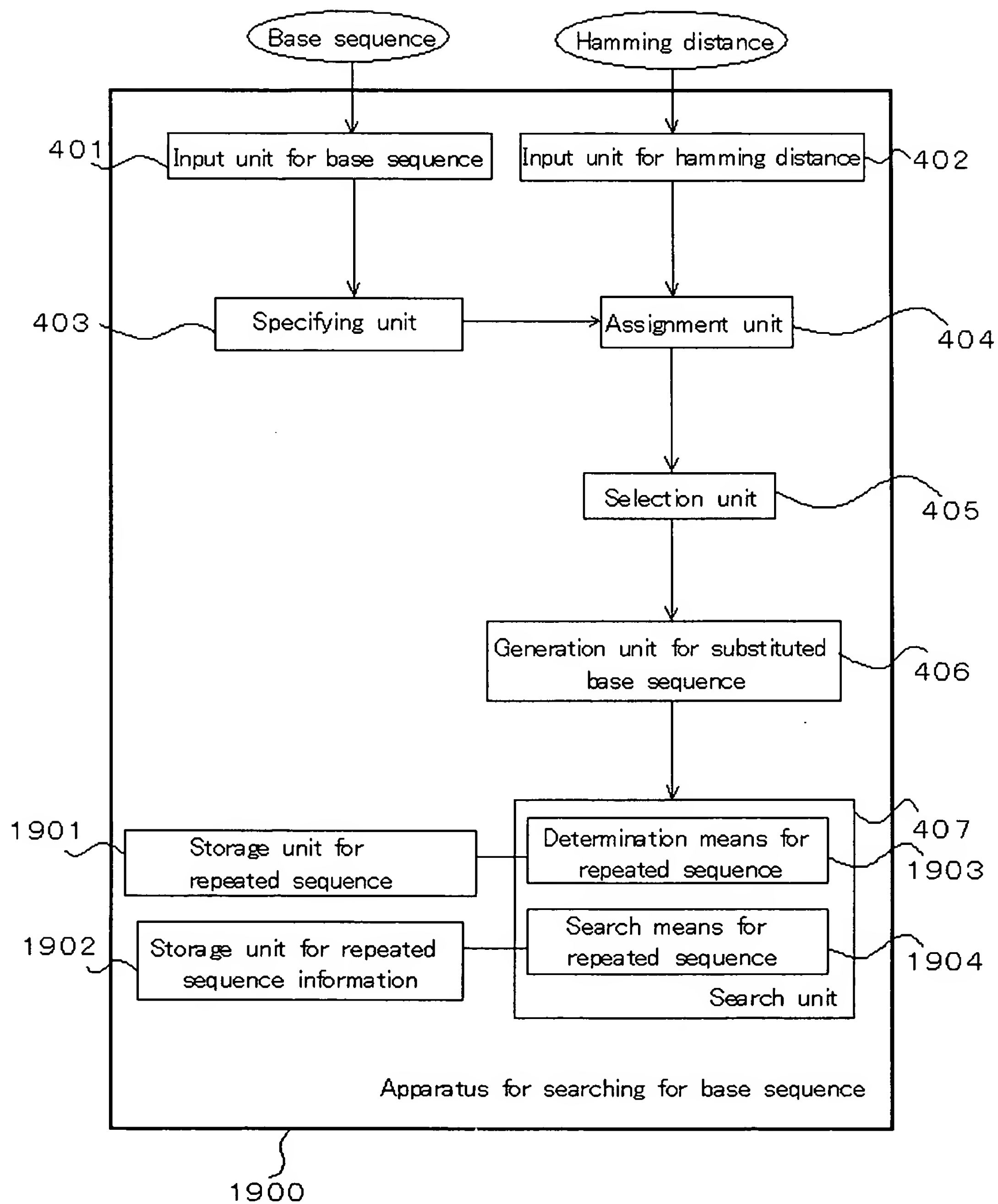


Fig 20

Identifier of repeated sequence	Repeated sequence
1	ACGGGUC...
2	GGGGGAAAA...
⋮	⋮

Fig. 21

Identifier of repeated sequence	Position of appearance
1	58210
1	37703
1	27503
1	30516
⋮	⋮
2	167
2	27367
2	112109
⋮	⋮

Fig. 22

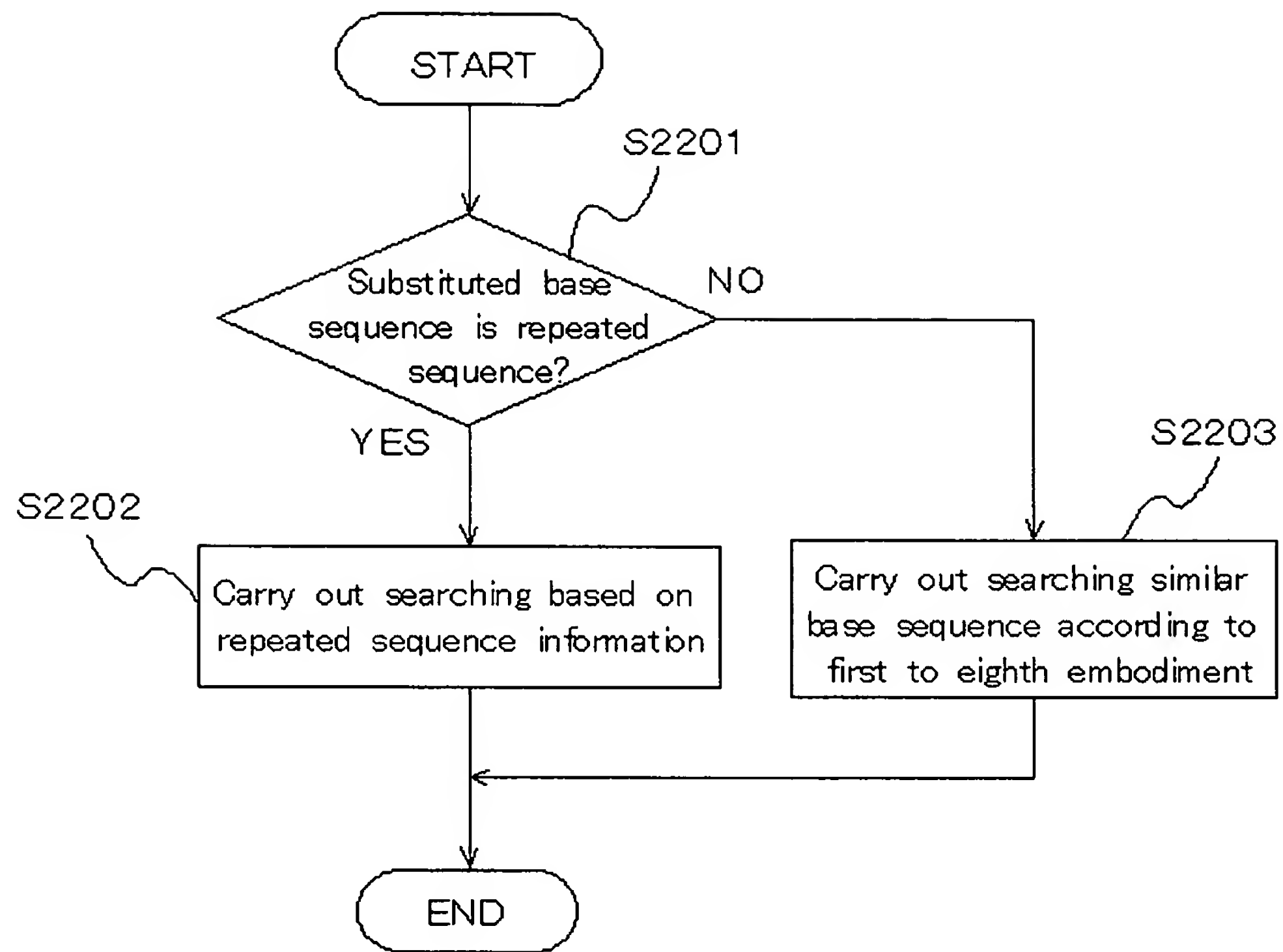


Fig. 23

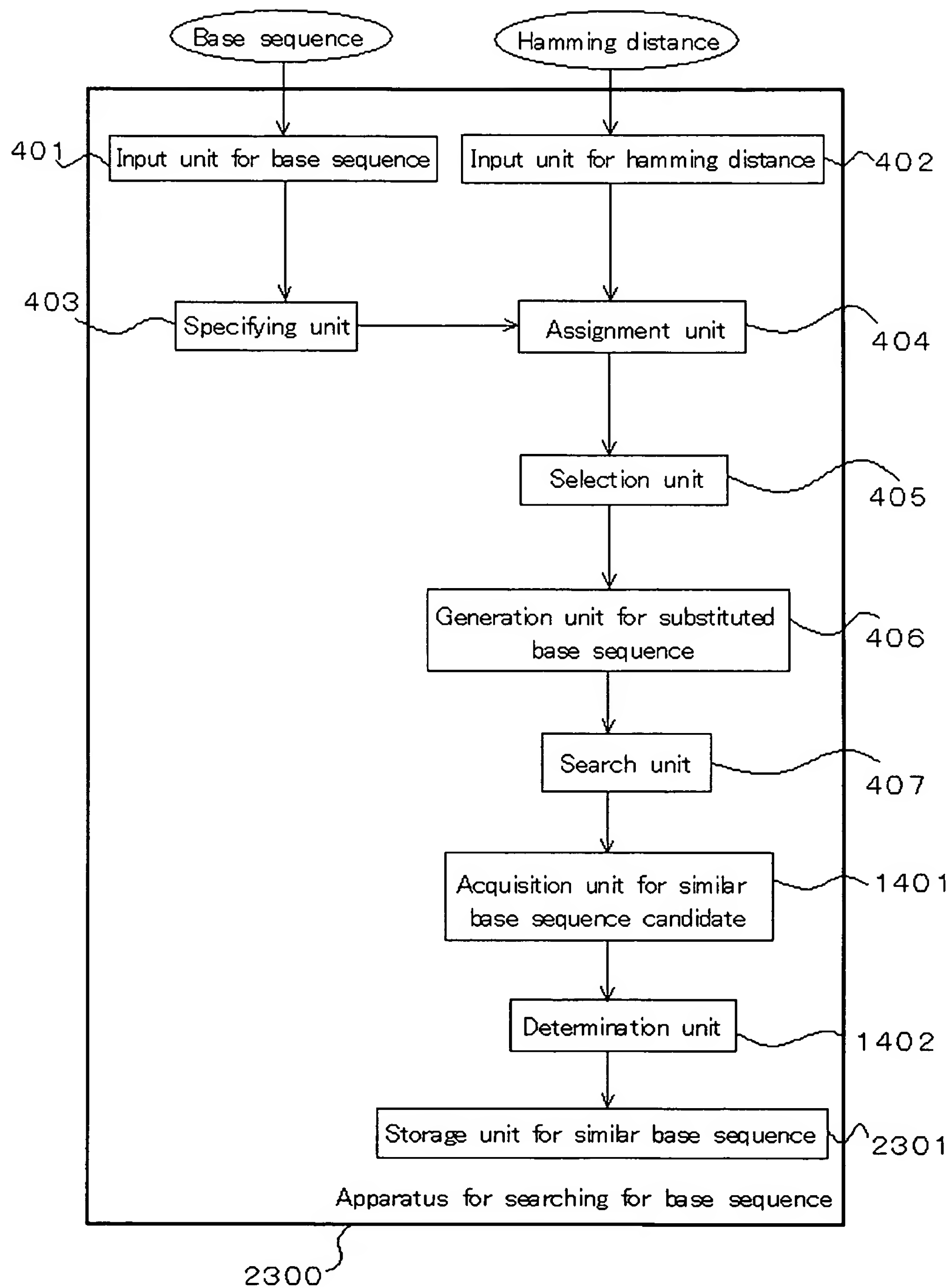


Fig. 24

Input for base sequence	Hamming distance	Similar base sequence
⋮	⋮	⋮

Fig. 25

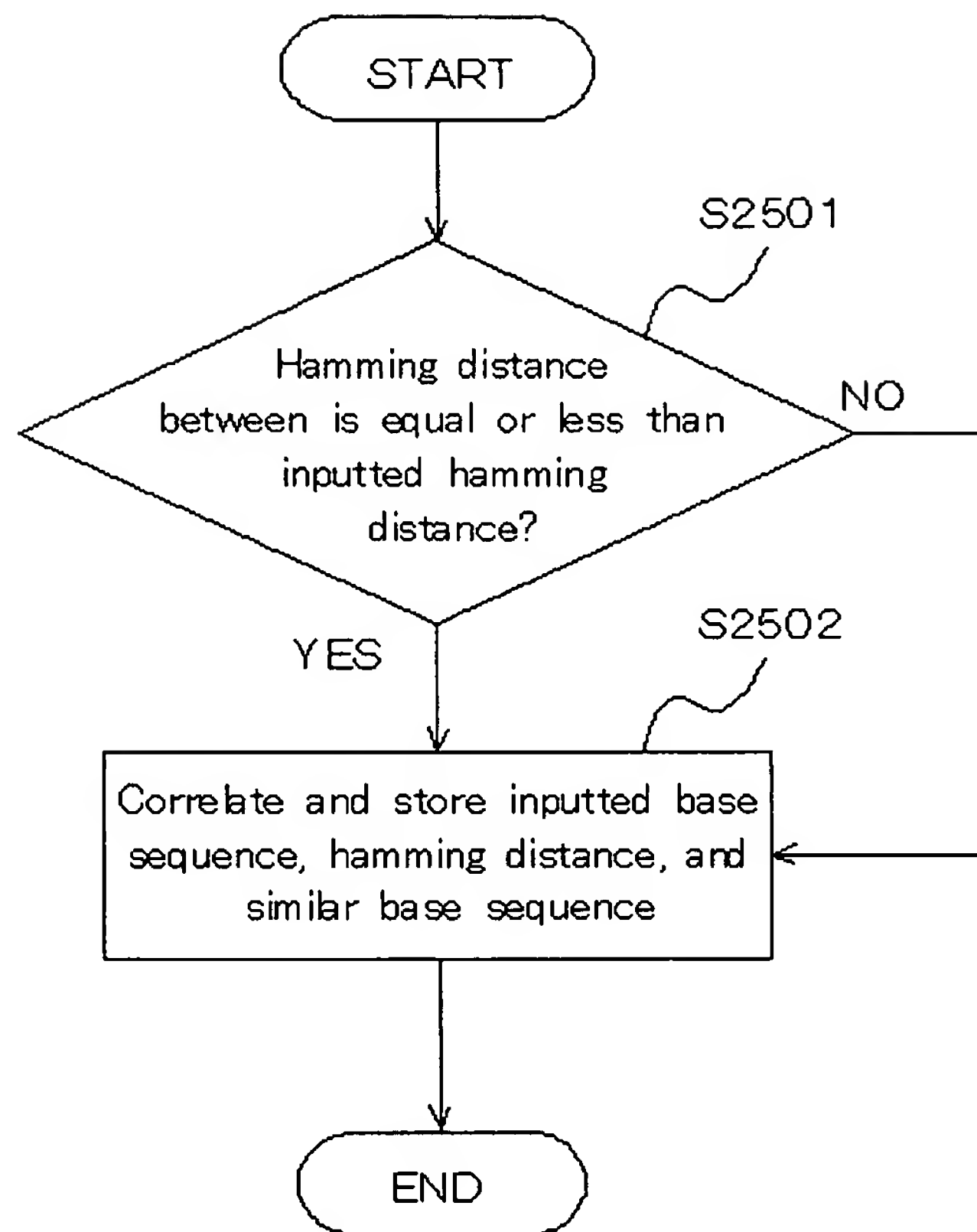


Fig 26

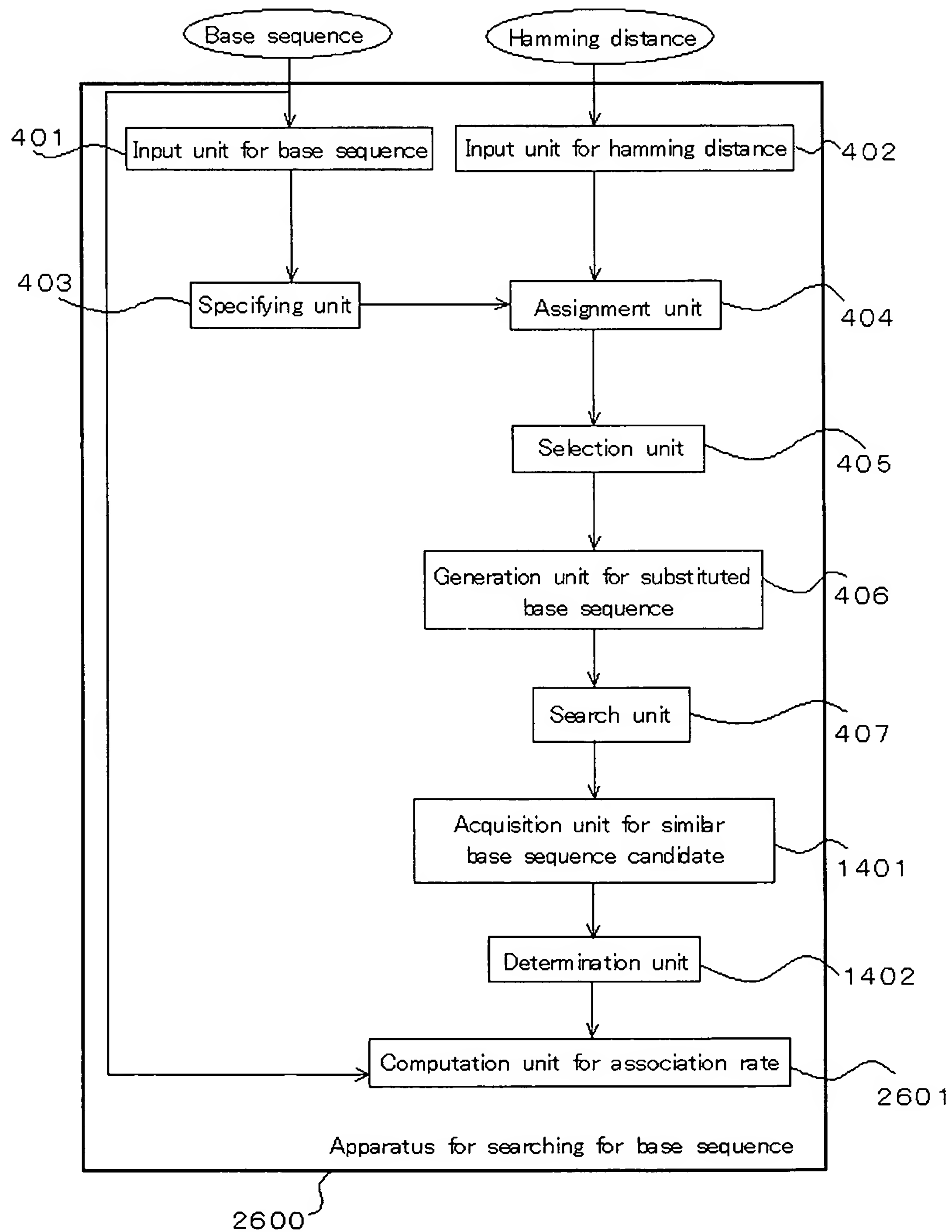


Fig. 27

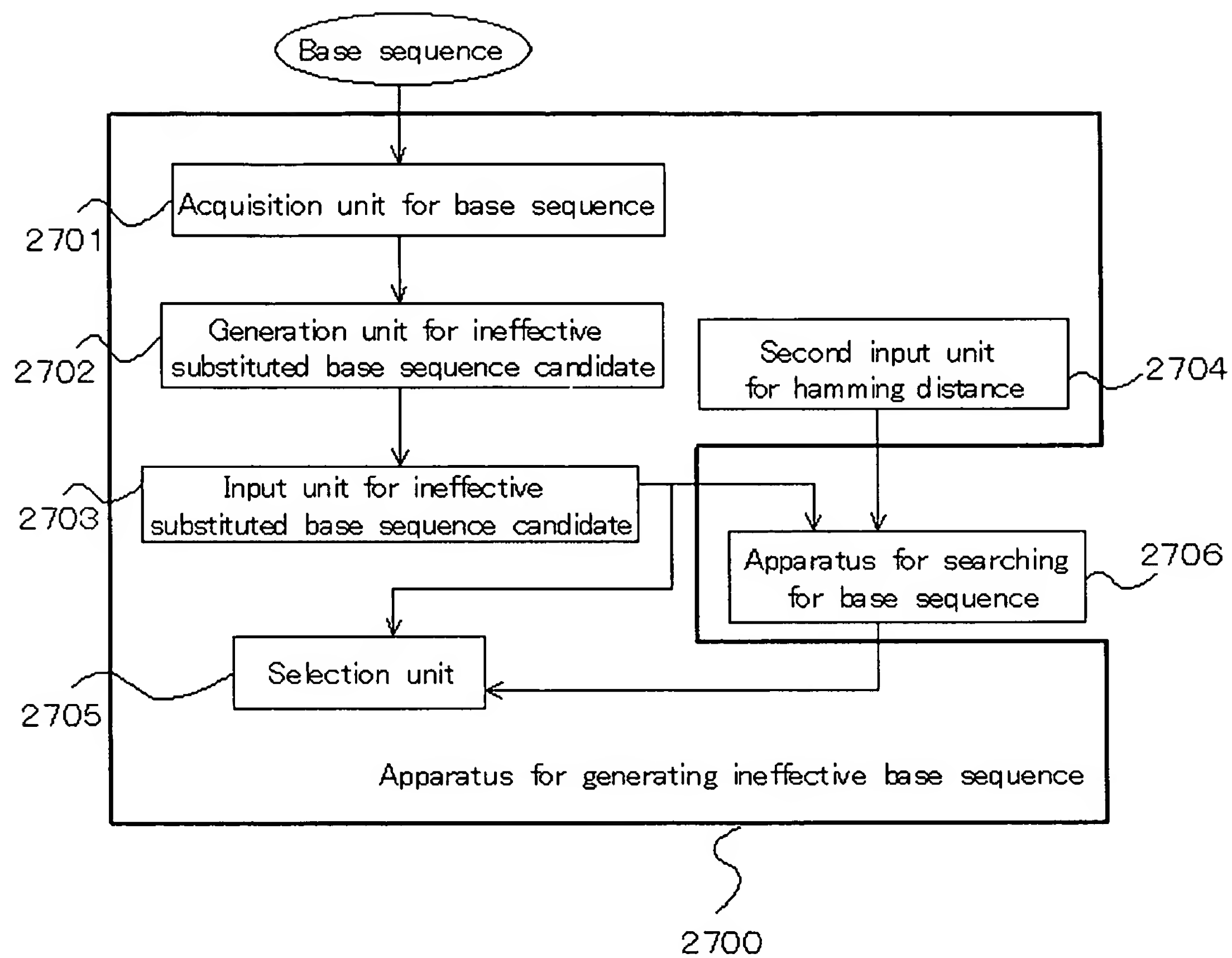


Fig. 28

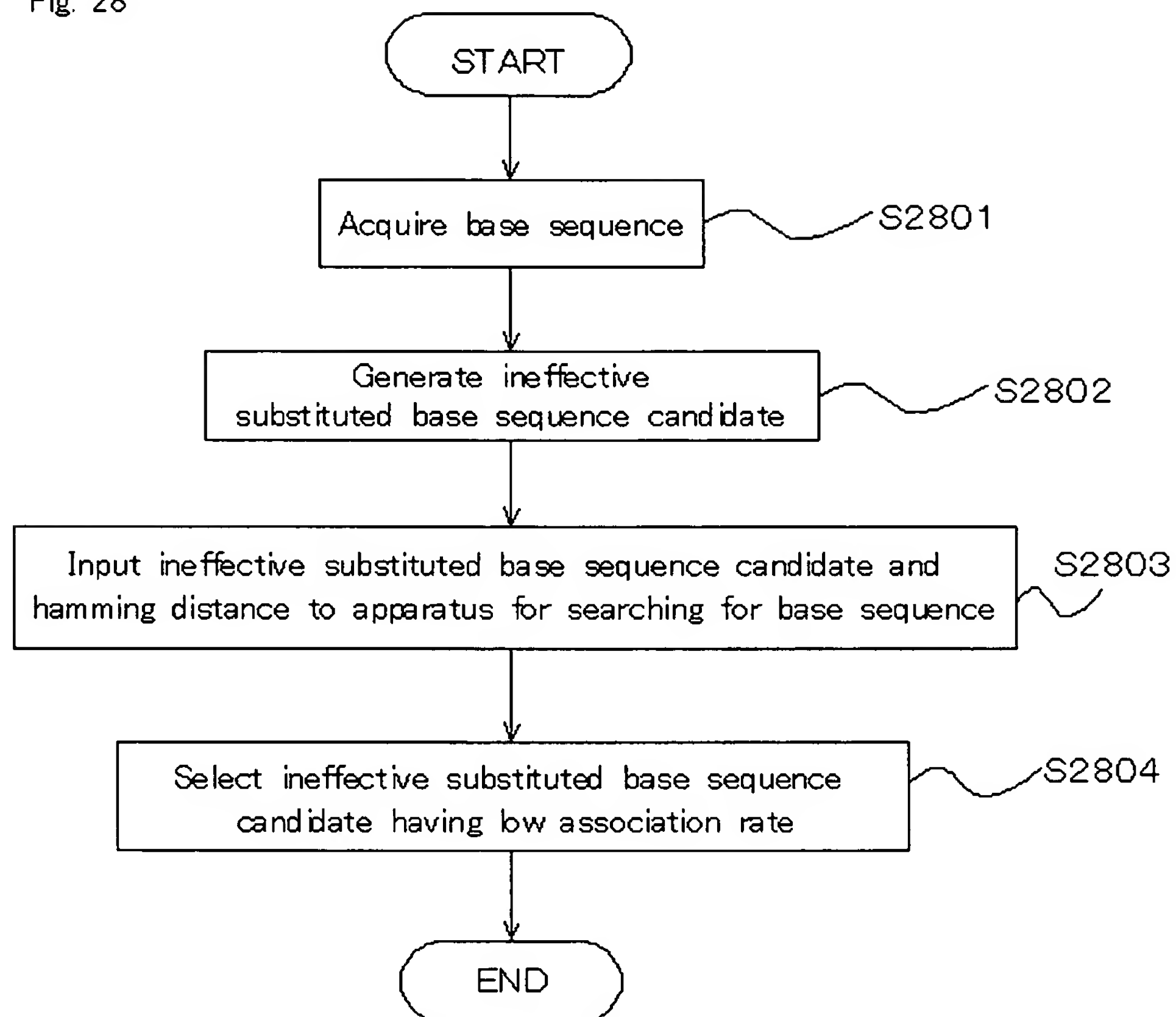


Fig. 29

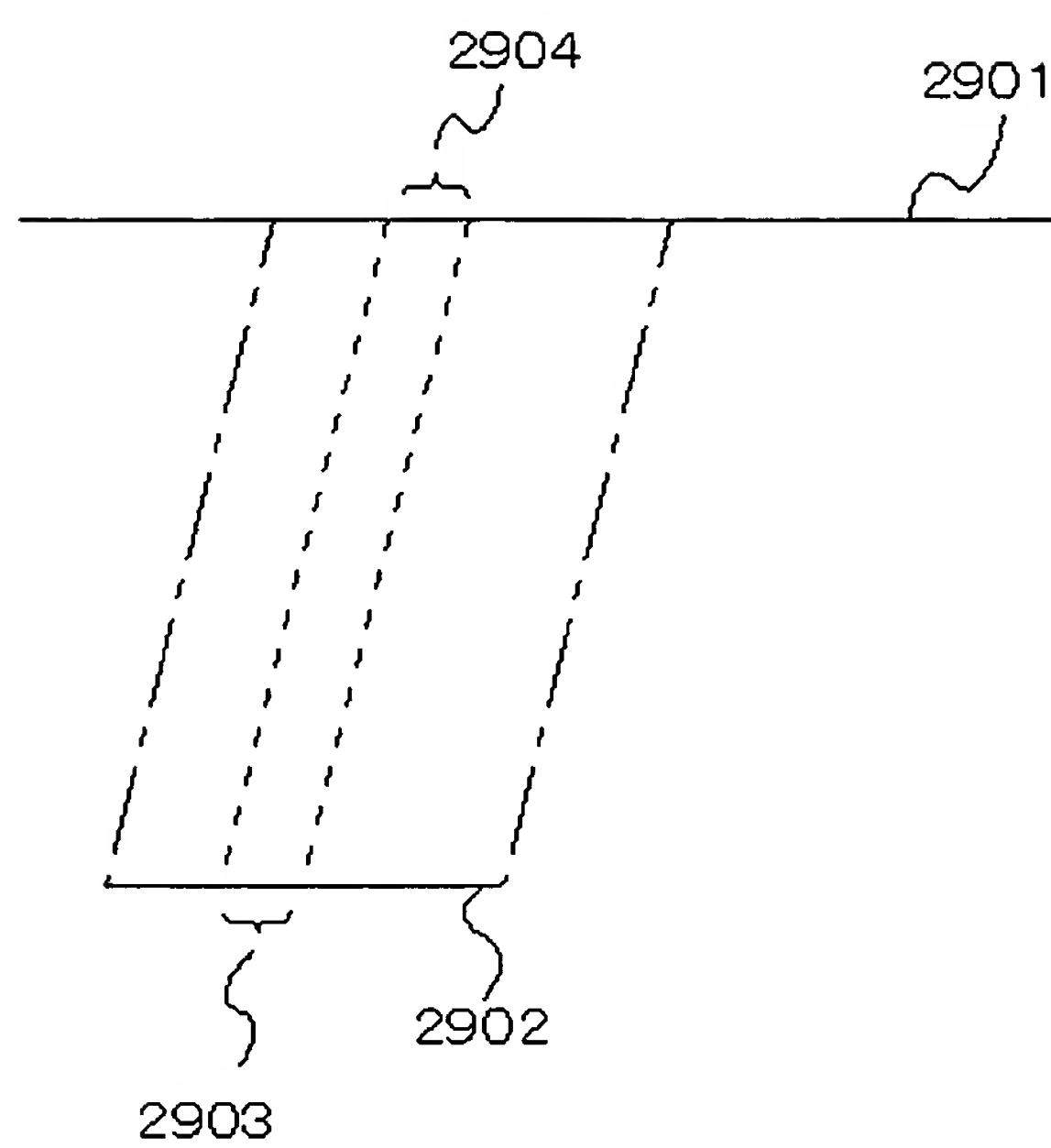


Fig. 30

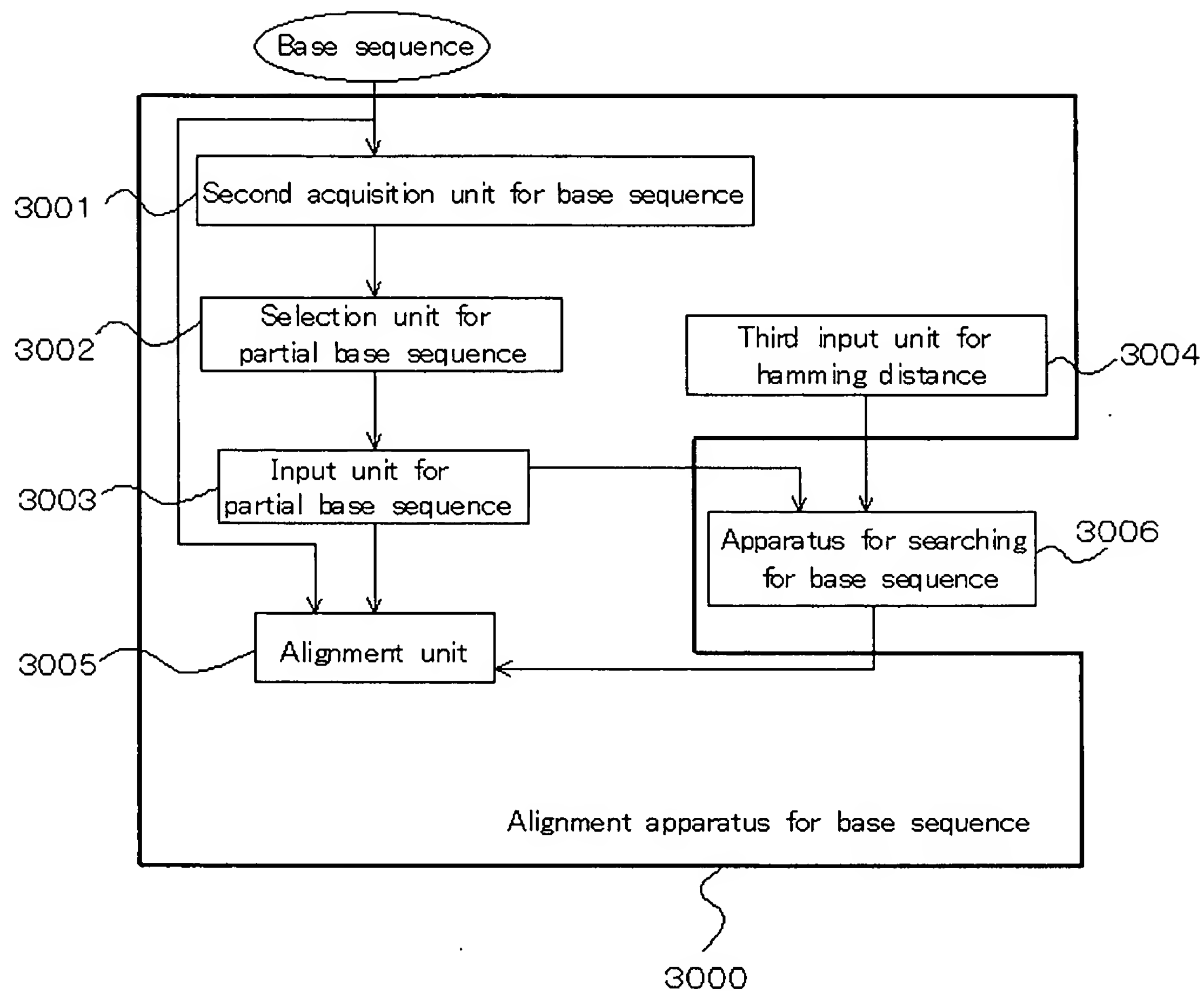


Fig. 31

